Assignment 4 - Hailey Pong

Problem 1

A) $\Delta w13 = d3$, $\Delta w23 = -d3$, $\Delta w14 = d4$, and $\Delta w24 = -d4$

B) output and error: a5 = 0.317, Error(W) = 0.467

Gradient and updated value:

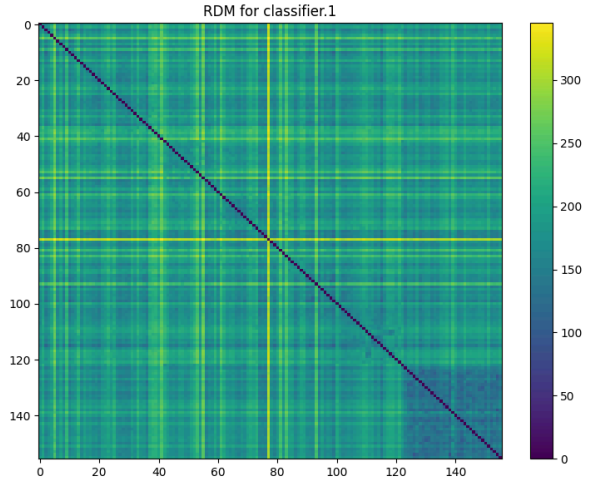
- w03: -0.058, -0.971
- w13: -0.058, 1.029
- w23: 0.058, -1.029
- w04: -0.074, 2.037
- w14: -0.074, -0.963
- w24: 0.074, 0.963
- w05: -0.296, -1.852
- w35: -0.216, 1.108
- w45: -0.148, 1.074

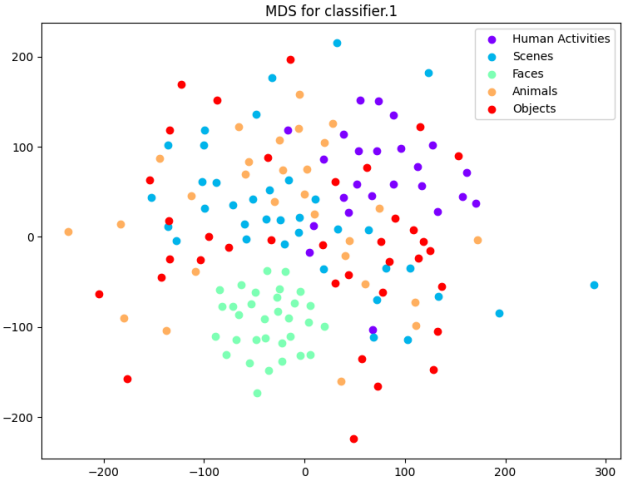
new output and error after the update: a5 = 0.388, Error(W) = 0.375

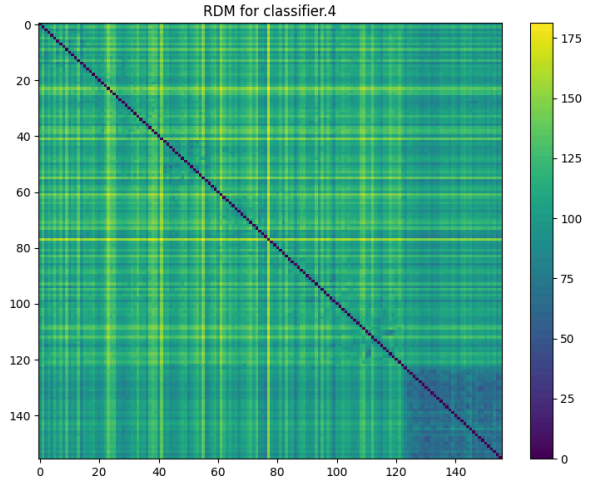
C)

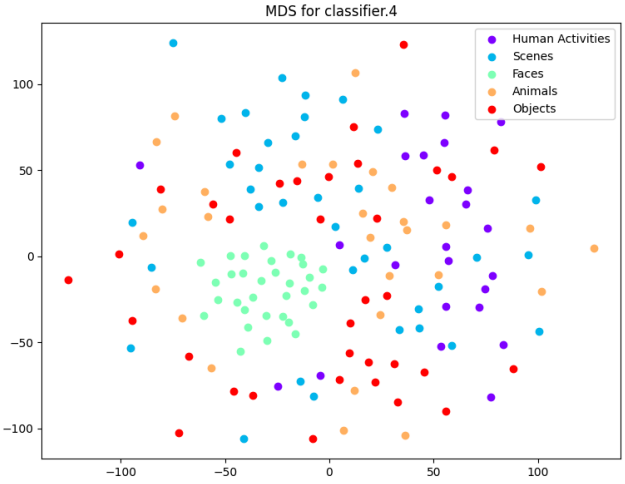
- output and error: a5 = −0.845, Error(W) = 3.404
- gradient and updated value for w03: -0.443, -0.778
- gradient and updated value for w13: -0.443, 1.222
- gradient and updated value for w23: 0.443, -1.222
- gradient and updated value for w04: -1.055, 2.528
- gradient and updated value for w14: -1.055, -0.472
- gradient and updated value for w24: 1.055, 0.472
- gradient and updated value for w05: -1.055, -1.472
- gradient and updated value for w35: -0.804, 1.402
- gradient and updated value for w45: 0.000, 1.000
- new output and error after the update: a5 = 0.636, Error(W) = 0.132

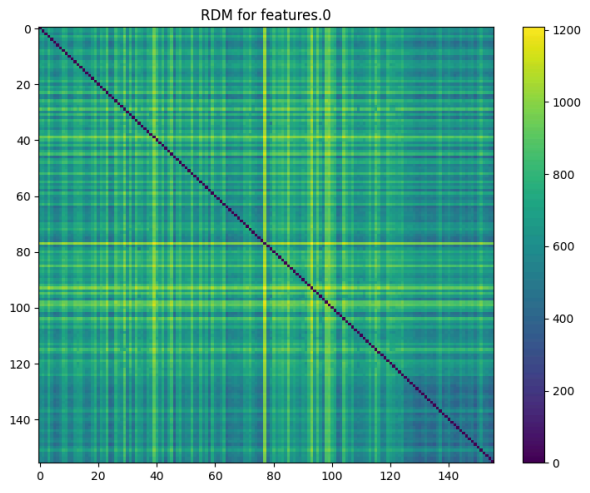
Problem 2)

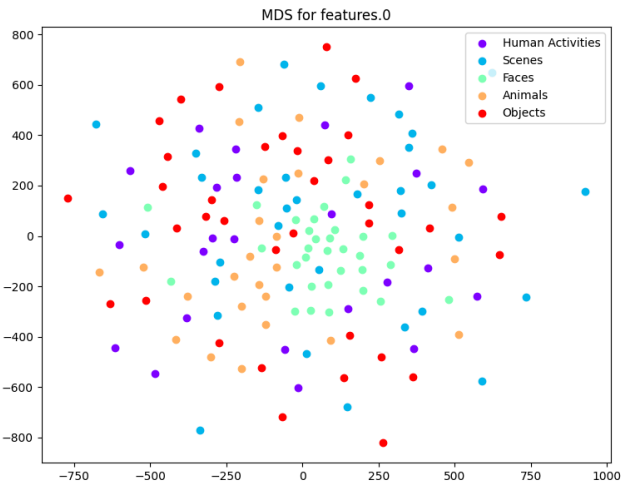

RDM for classifier.1

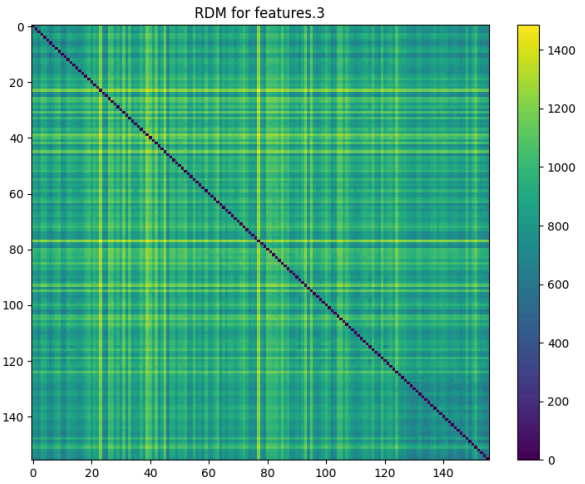

MDS for classifier.1
Human Activities
Scenes
Faces
Animals
Objects

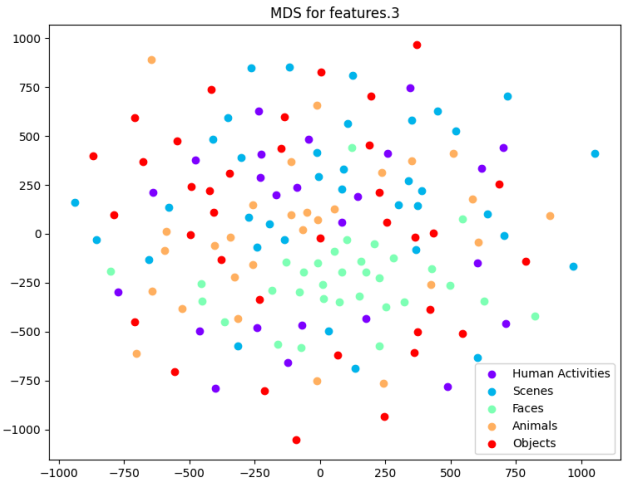

RDM for classifier.4

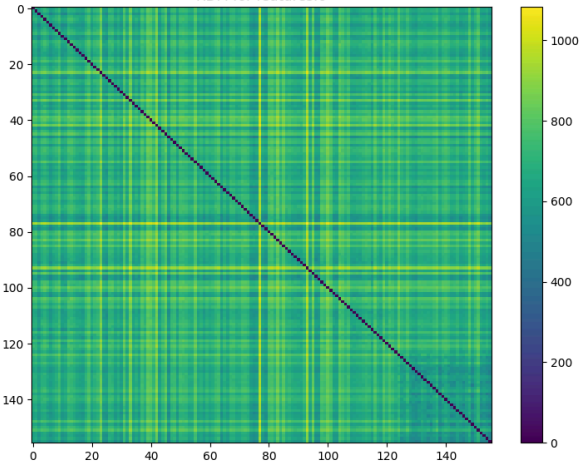

MDS for classifier.4
Human Activities
Scenes
Faces
Animals
Objects

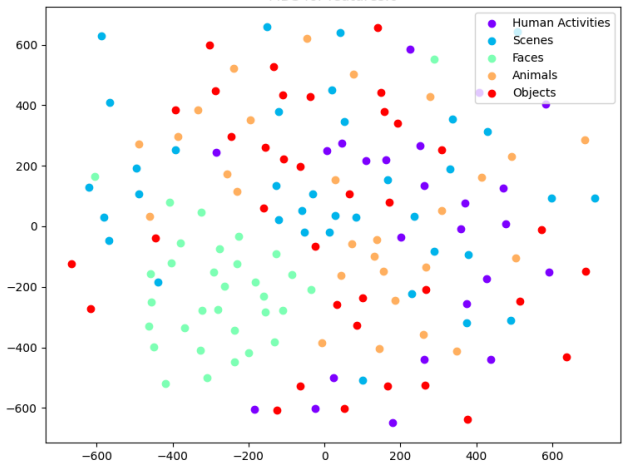

RDM for features.0

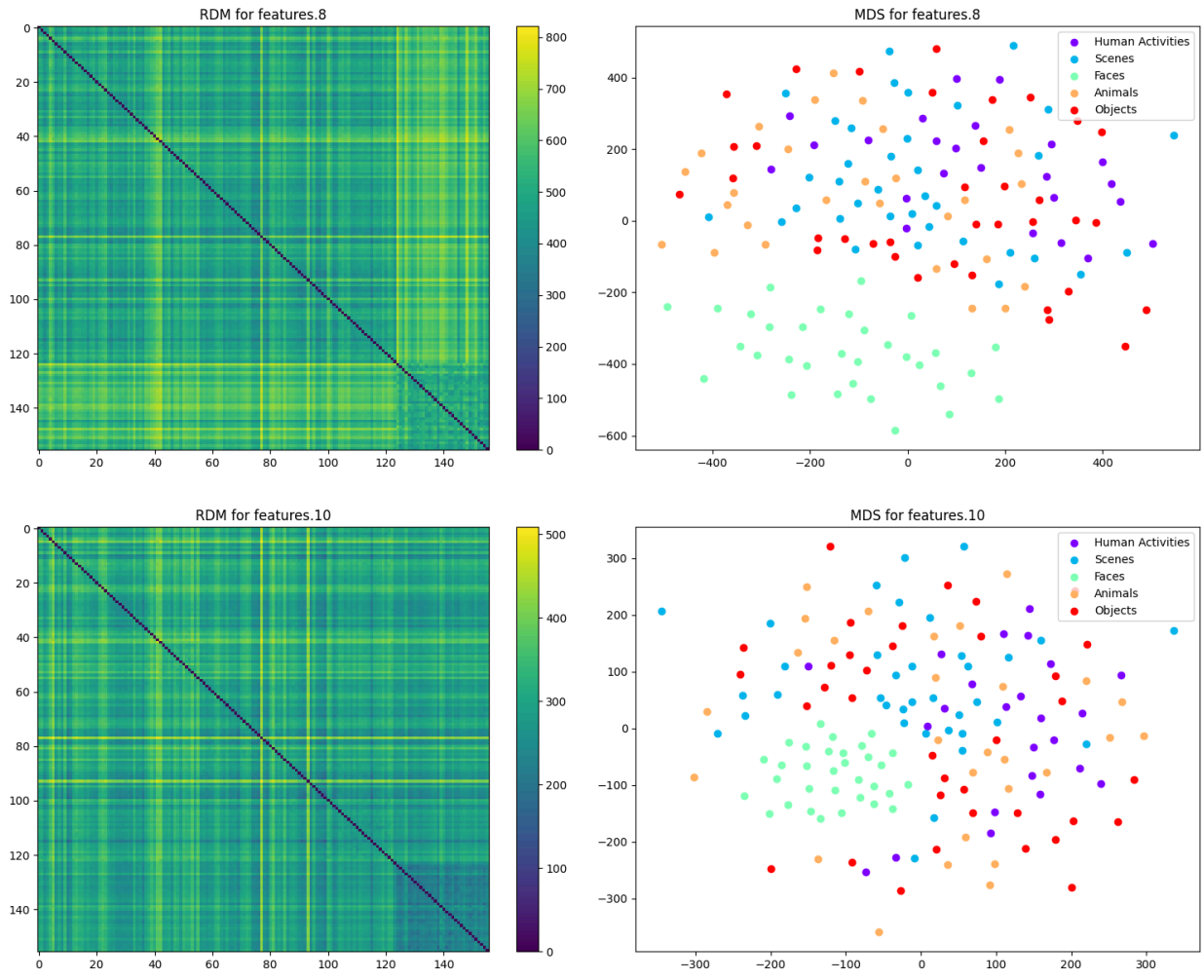

MDS for features.0
Human Activities
Scenes
Faces
Animals
Objects


RDM for features.3


MDS for features.3
Human Activities
Scenes
Faces
Animals
Objects


RDM for features.6


MDS for features.6
Human Activities
Scenes
Faces
Animals
Objects



The Representational Dissimilarity Matrix (RDM) helps visualize how a neural network processes and differentiates images. The matrix's symmetry shows that the dissimilarity between images A and B is equal to that between B and A, while its diagonal shows the dissimilarity between an image and itself, which is always 0. With color encoding emphasizing the degree of dissimilarity—darker regions indicate stronger similarity, and lighter regions reveal greater dissimilarity—clusters in the RDM imply that the network organizes images with comparable features.

Multidimensional Scaling (MDS) is a technique that reduces high-dimensional data to a 2D space, allowing us to visualize the relationships between image activations. In MDS, similar images tend to cluster together, with their proximity indicating how alike their features are. Distinct clusters suggest the network can differentiate between categories, such as faces, while overlapping clusters may indicate difficulties in distinguishing between certain image types. The separation or overlap of clusters reveals how effectively the model discriminates between categories.